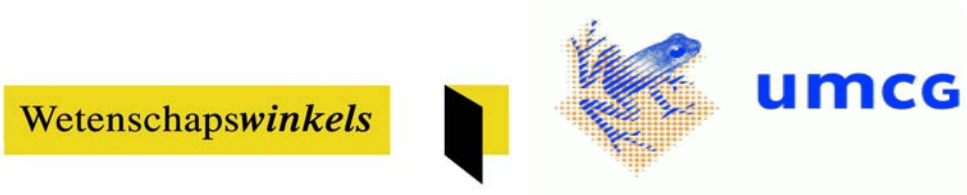
Self and parent administered joint assessment in patients with juvenile idiopathic arthritis

Jolanda G. Kaak, [No Value]



Self and parent administered joint assessment in patients with juvenile idiopathic arthritis

Supervisors:

W. Armbrust,
Department of Paediatric Rheumatology UMCG

J. Bouma & E. van Sonderen,
Faculteit Medische Wetenschappen RuG

Beatrix children's hospital UMCG Groningen
May 2010 – October 2010

J.G. Kaak

Colofon:

Vraagnummer:	GV 23/2010/A
Titel:	Self and parent administered joint assessment in patients with juvenile idiopathic arthritis.
Auteur:	Jolanda Kaak
Begeleiding:	Drs. W. Armbrust, kinderarts-reumatoloog Beatrix Kinderkliniek UMCG Dr. J. Bouma, Wetenschapswinkel Geneeskunde en Volksgezondheid UMCG Dr. F.L.P. van Sonderen, onderzoeker, UMCG
Vraagindiener:	Drs. W. Armbrust, Beatrix Kinderkliniek UMCG
Uitgave:	Wetenschapswinkel Geneeskunde en Volksgezondheid UMCG
Bezoekadres:	Hanzeplein1, De Brug, 9713 GZ Groningen.
Telefoon:	050 – 363 9080
Email:	j.tuinstra@med.umcg.nl
Datum:	Oktober 2010

Abstract
Objective. Important in the evaluation of disease activity in patients with juvenile idiopathic arthritis is (JIA) is the presence or absence of arthritis. Whether patients can accurately evaluate joints has never been researched. This study evaluated 1) the validity of self and parental assessment of arthritis in patients with JIA, 2) whether individual joints or joint groups are self assessed differently and 3) influence of demographic factors such as age, gender, condition, VAS pain and CHAQ scores in self assessment.
Methods. Each patient and accompanying parent was instructed to mark joints on a mannequin with different colours. Red for arthritis, yellow for doubt and green for no arthritis. Subsequently, a paediatric rheumatologist also marked joints on the mannequin. During analyses the judgement of the physician was considered the gold standard. Background variables were recorded for all patients and analyzed by descriptive statistics and tested with Chi-square and ANOVA.
Results One hundred and thirteen patients and their parents were included. One hundred eleven patients came for routine follow up. Two patients visited on own indication. In forty-three patients inflammation of at least one joint was found by the physician. Two patients (1.8%) and one parent (0.9%) did not signal the arthritis (i.e. false negative). Forty-three patients (38.0%) and forty-two parents (36.3%) signalled arthritis while it was not there (i.e. false positive). In one hundred and one cases (89.4%) there was agreement between parent and patient. Combining patient's and parent's assessments led to more false positive cases (n=48; 42.4%)) and only one less false negative assessment, when compared to the patient. Evaluation of assessments of individual joints showed that specificity and sensitivity ranged between 61.7%-96.4% and 33%-100%, respectively. In the assessment of the 4 extremities no actual pattern was found. However specificity and sensitivity are usually higher on the right side, with the exception of specificity in the arms.
By studying the demographic factors we found that, though females make more false positive assessments than males, difference between male and female was not significant over the four categories. The difference in VAS pain scores, CHAQ scores was significant. All patients with a high CHAQ (>1) or high VAS (>30) score signal arthritis, in half of the cases this is false positive.
Conclusions: Both patients and parents are not capable of a valid joint assessment. Underrating is low but overrating is high. Agreement between patient and parent is high. Based on this study we cannot state that assessment of one joint is better than the other. However assessment by joint shows wide ranges in both specificity and sensitivity. High VAS and CHAQ scores are predictive factors of patients signaling arthritis, however, in half of those cases there is no arthritis. To improve the validity of self or parent administered joint assessment an educational program needs to be developed.

Samenvatting (Dutch Summary)
Doel: Belangrijk in de evaluatie van ziekteactiviteit van patiënten met juveniele idiopathische artritis is aan- en afwezigheid van artritis. Of patiënten en ouders gewrichten kunnen beoordelen is nog nooit onderzocht. Deze studie evalueerde 1) de validiteit van zelfbeoordeling en beoordeling van ouders van gewrichten bij patiënten bekend met JIA, 2) Of individuele gewrichten en of gewrichtsgroepen beter of slechter beoordeeld werden, 3) en tot slotte de invloed van factoren zoals leeftijd, geslacht, ziekteduur, VAS-pijn en CHAQ score.
Methode: Patiënt en ouder werden geïnstrueerd afzonderlijk van elkaar op een mannequin alle gewrichten een kleur te geven. Rood voor artritis, geel voor twijfel en groen voor geen artritis. Vervolgens door de kinderreumatoloog, op eenzelfde mannequin alle gewrichten kleuren. Deze beoordeling werd beschouwd als de gouden standaard. Achtergrondvariabelen werden verzameld en geanalyseerd.
Resultaten: Honderddertien patiënten en hun ouder(s) werden geïncludeerd. Honderdenelf patiënten kwamen op reguliere controle, twee patiënten melden zich vervroegd. Er werd bij drieënveertig patiënten (38.0%) in tenminste artritis geconstateerd door de arts. Van deze drieënveertig merkten twee kinderen (1.8%) en daarvan 1 ouder (0.9%) de artritis niet op. Overschatting bleek groot. Drieënveertig patiënten (38.0%) en eenenveertig ouders (36.2%) signaleerden artritis terwijl het er niet was. In honderdeen gevallen (89.4%) was er overeenstemming tussen ouder en kind. Bij het beoordelen van de gewrichten afzonderlijk van elkaar bleek een grote variatie in sensitiviteit en specificiteit, respectievelijk 61.7%-96.4% en 33%-100%. Uit de beoordelingen per extremiteit bleek dat hier geen evident patroon in zit, maar het lijkt erop dat rechterzijde beter wordt beoordeeld dan de linkerzijde. Uit bestuderen van de demografische factoren bleek dat er meer meisjes fout-positief scoorden dan jongens, echter verschil tussen jongens en meisjes was niet significant. Wel is een significant verschil tussen CHAQ en VAS scores en de beoordeling per categorie gevonden. Alle kinderen met een hoge CHAQ- en VAS-pijnscore signaleren artritis, in slechts de helft van de gevallen is dit terecht.
Conclusies: Zowel patiënt als ouder blijkt niet in staat tot een valide beoordeling van de gewrichten. Er is meer overschatting dan onderschatting, hoewel dit voor slechts twee patiënten heeft geleid tot een vervroegd polikliniekbezoek. Overeenstemming tussen ouder en kind is hoog. Beoordeling per gewricht, door het kind, liet eenzelfde beeld zien.
Hoge CHAQ- en VASpijn-scores zijn voorspellend in het, terecht of onterecht, signaleren van artritis. Om geldigheid van zelfbeoordeling of beoordeling door ouder moet een educatieprogramma ontwikkeld worden.

Index

1 Introduction 5

1.1 Juvenile idiopathic arthritis
1.1.2 Epidemiology
1.1.3 Treatment
1.1.4 Prognosis and outcome
1.1.5 Remission

1.2 Assessment of disease activity in rheumatic diseases
1.2.1 Quality of life
1.2.2 Pain
1.2.3 Functional ability
1.2.4 Parent-child discordance
1.2.5 Standardizing joint assessment
1.2.6 Self assessment of arthritis
1.2.7 Aim of this study – Self or parent administered joint assessment in JIA

2 Materials and Methods 10
2.1. Patient selection
2.2 Measures
2.3 Analysis of clinical data

3 Results 13
3.1 Patient characteristics
3.2 Patient analysis
3.3 Joint analysis
3.4 Predictors

4 Discussion 24

Afterword and acknowledgements 28

References 29

1 Introduction

1.1 Juvenile idiopathic arthritis

Juvenile idiopathic arthritis (JIA) is arthritis of unknown aetiology that begins before the 16th birthday and persists for at least 6 weeks.(1) Pathogenesis involves both auto-immune and genetic factors in varying combinations.(1,2) Based on clinical and laboratory features, JIA can be classified into 7 different categories provided by the International League of Associations for Rheumatology (ILAR).(1)

Oligoarthritis: Arthritis in one to four joints at onset. 2 subcategories are recognised: persistent, in which the disease is confined to four or fewer joints and extended, in which arthritis extends to more than four joints after 6 months of disease. Oligoarthritis is the largest group, containing up to half of the patients.(1,3)

Reumafactor positive polyarthritis: Arthritis in 5 or more joints during the first 6 months. Rheumafactor (RF) is positive. This group is considered the same disease as adult RF-positive rheumatoid arthritis (RA) and contains 5-10% of the JIA patients.(4)

Rheumafactor negative polyarthritis: Arthritis in 5 or more joints during the first 6 months. 30% of the patients fall in this category. RF is negative.

Systemic JIA: Arthritis with, or preceded by a quotidian fever >39 degrees Celsius for at least two weeks, accompanied by at least one of the following: an evanescent rash, serositis, lymphadenopathy, or hepatosplenomegaly. 4-17% of JIA patients fall under this category.(3)

Enthesitis related JIA: Arthritis and/or enthesitis with at least two of the following: 1) sacroiliac joint tenderness, or inflammatory lumbosacral pain; 2) HLA-B27 positive; 3) First degree relative with medical confirmed HLA-B27 associated disease; 4) anterior iridocyclitis, usually symptomatic; and 5) onset of arthritis in a male >6 years of age. 10%

Psoriasis related arthritis: Arthritis and psoriasis, or arthritis and at least two of the following 1) dactylitis, 2) nail abnormalities, or 3) family history of psoriasis in a first degree relative. Containing less than 5% of the JIA patients, this is the smallest group.

Undifferentiated arthritis: Arthritis that fulfils criteria in no category or fit into more than one Up to 20% of children with chronic arthritis cannot be classified by the above criteria.(3)

The aim of this classification was to enable identification of groups of children with chronic arthritis, to help with research on pathogenesis, epidemiology, outcome studies and to compare and assess therapeutic trials.(3,4)

1.1.2 Epidemiology

JIA is the most common chronic rheumatic disease in children and an important cause of short-term and long-term disability. Prevalence varies between 15-160 per 100.000.

JIA is more common in girls than in boys, exception is enthesitis related arthritis which affects more boys than girls. (3)

1.1.3 Treatment

Although the primary goal of treatment of JIA is permanent remission of disease, no medication or combination of medications has been demonstrated to be effective in the majority of patients.(5)

Initial treatment in most patients with arthritis is a non-steroid anti-inflammatory drug (NSAID), in oligoarthritis usually combined with intra-articular injections. In polyarthritis Disease Modifying Drug's (DMARD) are required to achieve remission, sometimes in combination with a short period of low dose oral prednisone. First choice in DMARD is still methotrexate. If, within 6 months no remission is achieved, TNF-inhibitors are indicated.(2,3,6) In the last few years we have seen a shift toward these early aggressive therapeutic interventions.(6) In adults with recent-onset RA, initial combination therapy

(MTX and prednisone or MTX and infliximab) has shown a significantly faster improvement and a reduction of erosion.(7) In children a similar study will be started. In the treatment of JIA physical and occupational therapy also plays a critical adjunctive role.(5) Social workers and psychologists are very important providers of help with coping of JIA-patients and their families.(8,9)

1.1.4 Prognosis and outcome

The clinical course of JIA is variable, with unpredictable periods of arthritis and clinical remission. Its prognosis is difficult to predict. (10,11) In about 50% of the patients the disease follows a benign and self-limiting course, whereas in others the disease process is severe and unremitting and results in progressive joint destruction and serious disability.(3) Best predictors of a poor outcome are greater severity/extension of arthritis at onset, symmetric disease, precocious hip/wrist involvement, the presence of rheumatoid factor, prolonged active disease and early radiographic changes.(10) Patients with oligoarthritis generally have the best outcome, remission is achieved in 60-70%. Patients with RF-positive polyarthritis show the worst outcome, with a clinical course characterized, as in adults, by progressive and diffuse joint involvement, with growth disturbances (intra- and extra-articular) and high risk at developing osteoporosis.(12,13)
There has been a decline in the frequency of patients with severe physical disability over the years, but the proportion of patients who enter adulthood with active disease does not seem to diminish.(11) Nearly half of the patients with JIA had recurrent or persistent disease activity entering adulthood, with active arthritis, ongoing joint destruction and decreased quality of life.(12,13) Several studies noted lower employment rates for patients with JIA compared with age-matched local control subjects, despite comparable, if not better, academic achievements in patients with JIA.(12)

1.1.5 Remission

The criteria for inactive disease are: absence of active arthritis, absence of fever, rash, serositis, splenomegaly, generalized lymphadenopathy attributable to JIA, absence of active uveitis and normal ESR or CRP level. Combined with a physician's global assessment of disease activity indicating clinical disease quiescence.(5) Inactive disease can be divided further into clinical remission on medication (a minimum of 6 continuous months of inactive disease while receiving medication) and clinical remission off medication (12 months of inactive disease while receiving no medication).(5)

1.2 Assessment of disease activity in rheumatic diseases

In assessing disease activity in rheumatic diseases no single 'gold standard' is available to evaluate clinical status. Two quantitative indices are widely used in clinical trials: first is the American College of Rheumatology (ACR) Core Data Set, which includes swollen joint count, tender joint count, physician's assessment of global status, acute phase reactant-ESR or CRP, functional status(measured by health assessment questionnaire (HAQ), pain (measured by visual analogue scale (VAS), patients estimate global status, a radiograph in studies over 1 year or longer, and second is the disease activity scored (DAS), which includes swollen joint count, tender joint count acute-phase reactant, and patient assessment of global status.(14) In children the ACR core set criteria are used, the same as in adults except for the radiograph studies which are not included in the pediatric score. Thus, essential in evaluating disease activity is the presence or absence of arthritis.(14,15)
In the Netherlands follow up frequencies for patients with JIA vary between 2 to 12 times per year, depending on clinical course and medication. The standard follow up frequency for JIA patients is every 3 months. In periods of remission patients visit twice a year, in periods of

active disease patients are seen up to once a month. Unavoidable is that some patients are seen too often, which is socially and financially undesirable. Others, however, could develop arthritis in between visits, which could eventually lead to the reported consequences.

Attempts to measure and compare assessments of disease status by parents, patients and physicians have been undertaken. Different domains of the disease severity, such as perceived quality of life, rating of pain scales and measurement of functional ability have been studied. (16-19) Considering the importance of these domains we will discuss them separately.

1.2.1 Quality of life

Children with JIA have a significantly lower health related quality of life (HRQL) than is reported for healthy children in the general community.(20) The impact of JIA and its treatment on physical, functional and social outcome is reflected in impaired patient health and perceived quality of life. Health in young adults with JIA has been repeatedly analyzed using quality of life measures. Two studies showed that children and adolescents with JIA judged their quality of life lower than that of age- and sex-matched control subjects. (12,20) When patients and parents' ratings of disease activity and HRQL are compared a wide variation in results is seen.(19) One study showed that even though patients in remission had HRQL scores similar to those of healthy controls, their parents/proxies reported more fatigue than controls.(21,22) Sawyer et al found that children reported significantly better physical, emotional and social functioning than was reported by their parents.(20) Another study showed that there seemed to be good agreement between patients and their parents concerning quality of life, except for fine motor function, and that agreement was higher among younger children and among those who had had longer disease duration.(16)

1.2.2 Pain

Joint pain is a common symptom of JIA. Although substantial research in pain experienced by adults with RA has been done, pain related to JIA has been understudied.(20,23) This may be due to the difficulties in measuring paediatric pain.(23-25) Pain in children is best understood as a multi-factorial concept in which pain is the result of somatosensory, behavioural and environmental factors, and should thus be assessed multi-dimensionally.(26) Recent studies suggest that pain is quite common in children with different forms of arthritis and may be under recognized and under treated.(20,18)

Accurate pain assessment is the foundation for effective pain management in children.(24) The Faces pain scale revised (FPS-R) is the type of self report measure preferred by most children, parents and nurses.(27) Age is a significant predictor of children's ability to use the scale, the ability improves with age.(27) The FPS-R is advised between the age-range of 4-12 years. The Visual Analogue Scale, (VAS) a pre-measured line where the ends of the line represent extreme limits of pain intensity, is advised to use between 2-17 years.(25,28,28) In studies concerning JIA both these methods are used, though with different outcomes.(29) According to Thastum et al. disease duration is a better predictor of pain experience than age, however this could be explained by the assumption that younger children are less able to separate pain intensity from pain discomfort.(23)

As to the agreement between parents, patients and physicians in assessing pain, results vary. However, they all showed that there is a moderate to common disagreement between parent and child. A predictor of disagreement is a child with depressive symptoms, which leads to difficulties in communication with their parents.(18) Agreement is better between mother and patient than between father and patient. Physician provided higher (i.e. worse) ratings for pain than children and parents, which led to the conclusion that there is a poor level of agreement

between physician and patient and a moderate level of agreement between parent and physician.(18,29)
Thastum et al investigated pain coping strategies and states that JIA patients may differ from healthy children with regard to experimental pain as well as to their use of pain coping strategies. He found that a group of healthy children was able to endure experimental pain longer than the group of patients with JIA.(23) Measuring the pain threshold (PT) endorses this, by showing that the PT's in JIA patients are significantly lower than in their healthy peers, both in children with active inflammation as well as in children without detectable inflammation. In the patients with active inflammation PT's are significantly lower than in the patients without detectable inflammation.(30)
Disease activity only partly explains the variation in the pain experience of patients with JIA, and it is possible that psychosocial factors contribute substantially to the child's pain perception.(23)

1.2.3 Functional ability

Functional ability is regarded as an important outcome measure in the care of children with JIA.(31) The Childhood health assessment questionnaire (CHAQ) is a disease specific health instrument that measures functional ability in daily living activities in different domains in children with JIA.(32) Many studies show a significant higher CHAQ score in children with JIA, compared to healthy subjects.(32)
Reports on the agreement between parents and physicians concerning ratings of functional ability vary, ranging from fair agreement to agreement in only 43% of the cases. This may be due to different ways of measuring. The most simplified method, according to the author, showed higher levels of agreement.(19) Palmisani also found that in cases of discordance, the parent overrated (i.e. a worse physical function) compared to the physician.
A worrying fact is that recent studies show that children with JIA are considerably less active than their peers (33) and are at risk of losing benefits of physical activity. (33,34) The effects of activity programs seem promising; they are safe and may result in an important improvement in physical function.(35)

1.2.4 Parent-child discordance

Factors that may contribute to differences in the assessment of parents and children concerning the perceived quality of life, pain and functional ability are for example inadequate communication between parents and adolescents and it is also possible that children have adapted better to their circumstances and perceive their HRQL better than their parents.(20) Achenbach studied the degree of consistency between different informants' (children, parents, proxies, teachers) reports of the behavioural/emotional problems of subjects aged 1 ½ to 19 years. He stated that children aged 6-11 have higher correlations with parents than adolescents.(36)

1.2.5 Standardizing joint assessment

In rheumatoid arthritis (RA) a clinical evaluation is considered important for the evaluation of an individual patient's disease in daily clinical practice as well as in clinical trials.(37) Systematic assessment of swelling and tenderness in joints, or joint counts (JC) has been cited as the most specific measure of disease activity in RA and has been shown to be predictive of mortality.(38) A variety of measures have been used in clinical research and clinical care. Part of the assessment is always joint count by a physician. In which a physician states if the joint is swollen, tender or has limited motion or deformity. Between physicians high correlations are found.(15) The number of joints that are assessed depend on the method: The Ritchie index includes 52 joints; the modified Ritchie index contains 42 joints. Recent studies in

adults prove that a 28 joint count is correlated similarly with other measures of clinical status and is as effective in clinical trials as are joint counts involving more joints.(39) This is not studied in children yet.
Methods for assessing arthritis have been extensively researched in RA. Based on the fact that clinical trials in RA must be supplemented by long-term observational studies to assess results of therapy in regard to long-term outcomes, the most simple and effective method of collecting these data is through patient self- report questionnaires.(14) Other research also states that it would be easier and less time consuming if a patient can produce a valid self-administered joint count.(39)

1.2.6 Self assessment of arthritis

Self reported outcomes in RA fulfil a central role in the measurement of response to treatment both in clinical trials and in routine practice.(40) Therefore multiple studies have assessed the reliability, validity and sensitivity of self-reported joint counts.(40) Methods for assessing joint tenderness and swelling are different. A text format or a pictorial format, in forms of a mannequin, or both, are used. A review that compared these two formats found that the mannequin format fared better than text.(40)
Some authors reported good reliability and suggested that patients self-reported joint counts can be used to measure disease activity in RA, while others found poor to moderate reliability.(39,41) A review of the studies assessing validity of self reported joint count stated that a tender joint count has moderate to marked correlation with those performed by a trained assessor. In contrast, swollen joint counts demonstrate lower levels of correlation.(40) This is also seen in a recent study comparing assessment of patient, physician, nurse and ultrasonography showed the same results, with underrating of physicians concerning tender joints and underrating of patients concerning swollen joints.(42) Improvement of the efficiency and quality of care by integrating a self reported tender joint count will be studied in the future.

1.2.7 Aim of this study – Self or parent administered joint assessment in JIA

In the assessment of disease activity of arthritis self assessment seems important. Self assessment of pain, quality of life and functional ability have all been researched, with different outcomes. In general we can state that parents overrate the different domains of disease status compared to their children.(16,18,19,29,31)
Self assessment of arthritis has been extensively researched in adults, but is missing in children. In adults this seems to be the route for the future. Not only because it will be less time consuming for physician if patients appear to be capable of a valid joint count, but it will also be valuable in clinical trials. If patients appear to be unable to assess their joints there is a role for education. There are many potential benefits of a self reported joint count. Involving patients in disease activity assessment may enhance self management behavior, and ultimately improve health outcomes. Self-management programs in arthritis have been shown to improve health status, reduce pain and fatigue, and increase self-efficacy.(43,44) Active engagement with one's chronic disease has been shown to be associated with health improvement.(40) Research also showed us that understanding the medical situation is beneficial for coping of parents and ultimately leads to improvement of functional ability of the patient.(8) It is important to study the validity of assessment of both parents and patients, because the majority of the patients are young at disease onset.
In the present study, we investigated the validity of self or parent administered joint assessment in children with JIA. Secondly we studied which joints are assessed best or worst. Finally we attempted to identify if factors such as gender, age, disease duration, VAS-score, CHAQ-score were relevant or significant predictors of patient-physician discordance.

2 Materials and Methods

2.1. Patient selection

The study group comprised all patients that were diagnosed with JIA according to the revised ILAR criteria and that visited the outpatient clinic of two tertiary pediatric rheumatology centers between May and July 2010. The centers include the Beatrix Children's Hospital (BKZ) in Groningen and the Wilhelmina Children's Hospital (WKZ) in Utrecht. Verbal informed consent was obtained from parents and child (when applicable).
Patients aged between 4 and 18 were included. They were accompanied by mother, father or both. Families in which the accompanying parent did not live with the child on a daily basis were excluded. Patients with insufficient knowledge of Dutch or English language are also excluded.

2.2 Measures

15-30 minutes prior to the visit to the pediatric rheumatologist, the child and parent(s) were asked to independently assess the disease activity at that time, by usage of a mannequin representing all joints, except for those too difficult to identify or assess (acromioclavicular joints and lumbal/thoracal spine). The format with mannequin and instructions was developed for this purpose, and based on the mannequin used by rheumatologists.
Patients and parents were instructed to mark the joints with different colors. Arthritis was marked red, doubt was marked yellow, and the joints in which there was no arthritis were green. We explained that joints in which the disease was active according to patient or parent needed to be marked with red. Other explanation, with respect to the term 'arthritis', like warmth or swelling, was not given. Instructions were given on paper and elucidated verbally, always by the same observer (JK). We emphasized that this study was about the current status of inflammation of joints, not about pain.
Immediately after the interview and rheumatologic physical examination the pediatric rheumatologist marked the joints on the same format of a mannequin, with red, yellow or green.

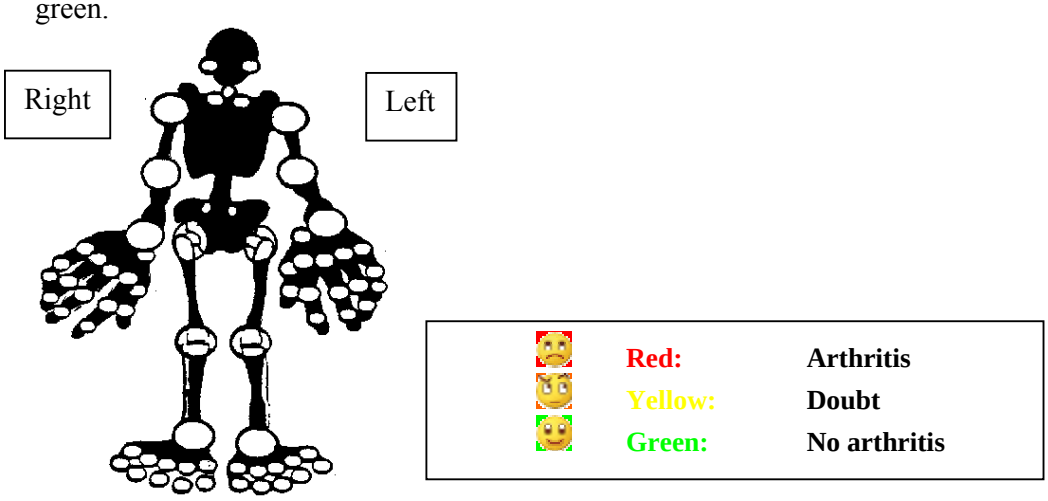


Figure1: Self-administered form to indicate arthritis.

Prior to the visit we assessed functional ability by asking patients to complete the validated Dutch translation of the Childhood Health Assessment Questionnaire (CHAQ). This includes questions for the following domains: dressing/grooming, arising, eating, hygiene, reach, grip and activity. Outcome ranges between 0 and 3, 0= able to do without difficulty; 1= able to with some difficulty; 2 = able to with much difficulty; 3 = unable to. (32)

When younger than 8 years, parents completed the form, with or without children's aid. Furthermore, we asked patients and parents to mark the level of pain in the last week on a 100m VAS pain scale, separate of the VAS pain and VAS severity in the CHAQ.

Clinical assessment The medical chart of each patient was reviewed for the following information: age, sex, condition, and disease duration at the study visit. Laboratory results (if performed) included C-reactive protein and erythrocyte sedimentation rate (ESR) and were linked to the patients.
Concerning the classification of the condition we categorized according to the ILAR criteria, but chose not to differentiate between rheumafactor positive and rheumafactor negative polyarthritis. Patients with extended oligoarthritis are also shared under the polyarthritis, based on outcome and clinical presentation at follow up.(5)

2.3 Analysis of clinical data
Outcomes were divided into two groups: green and not green, thus 'no arthritis' and 'doubt/ arthritis'. Doubt is considered to be 'arthritis' from this moment. Based on the fact that clinical assessment by a rheumatologist is the most important predictor of disease activity we considered the clinical assessment of the pediatric rheumatologist as the gold standard.(37) A combination of parent and patient was made in which the worst (i.e. yellow or red) assessment was registered.
The outcomes were divided into four different categories:
1. Correctly signaled that there is no arthritis (i.e. true negative)
2. Not signaled arthritis, while there is (i.e. false negative)
3. Signaled arthritis while there is not (i.e. false positive)
4. Correctly signaled arthritis (i.e. true positive)

We studied the outcomes on two different levels: analysis by patient and analysis by joint. The difference being that in the patient analysis every patient is labeled with a category instead of the joint analysis in which every joint or joint group is labeled with a category.

a Analysis by patient: in each individual the assessment of 69 joints was compared to that of the physician. Based on this, patients were categorized by the above categories.
At this level category 4 was split into 5 subcategories:
4.a. signaled arthritis at correct location
4.b signaled arthritis at wrong location
4.c. signaled arthritis both at the correct and wrong location
4.d. signaled arthritis at the correct location, and missed arthritis at other location(s)
4.e signaled arthritis, at correct and wrong location, and missed arthritis at other location(s)
For example: Patient judged: 30 joints correctly that there is arthritis;
in 2 joints no arthritis while there is; in 30 joints arthritis is without being there;
in 7 joints arthritis while it is there.
This patient is put in category 4, because of signaling arthritis. Because he also missed 2 joints with arthritis and he overrated 7 joints patient X falls in subcategory e.
Based on the first conclusions we focused further analysis on the assessment made by patients

b Analysis by joint. To answer the question which joints or which extremity is assessed best or worst joints were evaluated separately and in groups, by extremities.

We combined joints to form 4 extremities: Left arm and right arm, left leg and right leg. The arms consist of shoulder, elbow, wrist, carpometacarpal (CMC) joint, metacarpophalageal (MCP) joints, proximal interphalageal (PIP) joints and distal interfalangeal (DIP) joints. The legs are a combination of hip, knee, ankle, metatarsophalangeal (MTP) joints and interphalangeal (IP) joints.

If one or more of the joints in a group was marked with arthritis by patient or physician it was considered as arthritis in the concerning extremity. We also combined small joints to form left and right foot and left and right hand.

For example: patient localizes disability, stiffness or pain in his left arm and marks his left wrist, while the physician finds arthritis of the left elbow. This way arthritis is found in the left arm, both by patient and physician, though patient does not localize correct.

Statistics We used descriptive statistics to analyze differences between patients, parents and physicians. Physicians assessment is considered gold standard, thus sensitivity and specificity were calculated on patient and joint level.

To explore which patient characteristics (i.e. backgroundvariables) could influence the outcomes of self assessment, we used descriptive statistiscs. Differences of 10 % were considered clinically relevant. The Chi-square was used to assess the homogeneity of the categorical variables by category of assessment. Null-hypothesis was that there is no difference between gender/ agegroups/ condition/ disease duration/ CHAQ-score /VAS-pain score in Cat.1/ Cat.2/ Cat.3/ Cat.4 of assessment, One-way ANOVA was used to identify if mean age, disease duration, CHAQ-score and VAS-pain differed significantly by category of assessment. A P-value lower than 0.05 was considered significant in both tests.

For analyses of clinical data SPSS 16.0 for Windows was used.

3 Results

3.1 Patient characteristics

One hundred and fourteen patients were initially identified to participate. One patient was excluded due to an invalid assessment, (as judged by observer; patient's favorite color was red). None of the patients refused participation. Thus, one hundred and thirteen children were included. Patient characteristics are shown in Table 1. None of the patients in this group was diagnosed with psoriatic arthritis.

111 patients visited the clinic for routine follow-up. Two children visited on own request, because of signaling arthritis. One of these visits was valid, there was arthritis. The other marked arthritis in 37 out of 69 joints, all of them were false negative.

Table 1. Patient characteristics

Characteristic		*Sample (%) n = 113*
Gender	Male	37 (32.7)
	Female	76 (67.3)
Age	<9 years	27 (23.9)
	9-12	35 (31.0)
	>12	51 (45.1)
Condition	Oligoarthritis▪	43 (38.1)
	Polyarthritis *▪▪	55 (48.7)
	ERA**	4 (3.5)
	Systemic JIA	9 (8.0)
	Other arthritis***	2 (1.8)
Disease duration	=<12 months	22 (19.5)
	=> 13 months	91 (80.5)
CHAQ-score†	0	45 (39.8)
	=<1	45 (39.8)
	>1	19 (16.8)
	(missing)	(4 (3.5))
VAS-pain (child) ††	<30	85 (75.2)
	30-60	11 (9.7)
	>60	13 (11.5)
	(missing)	(4 (3.5))

▪ minus extended oligoarthritis
▪▪ plus extended oligoarthritis
*Rheumafactor positive and negative
** Enthesitis related arthritis
***Undifferentiated arthritis not otherwise specified
†Children's health assessment questionnaire. Minimum is 0; maximum is 3
††Visual analogue scale, measured in mm. Minimum 0; maximum 100

First we studied the assessment focusing on the patient. According to their assessments patients were labeled with one of four different categories, category 1 containing the patients that judged there was no arthritis while there was no arthritis (i.e.true negative); Category 2 containing the false negative-; category 3 containing the false positive-; category 4 containing the true positive assessments. In the fourth category a difference was made between localizing

arthritis at the correct or at the wrong location. This way we were able to calculate specificity and sensitivity of the complete joint-assessment by patient and parent.
Secondly we studied the joints individually and in groups, in which the assessments of joints were organized by the four different categories and again sensitivity and specificity were calculated.
Finally we studied the patient characteristics that could hypothetically be of predictive value in self assessment of joints.

3.2 Patient analysis

Arthritis of at least one joint was established by the physician in 43 (38.0%) out of 113 patients. 84 patients (74,3%) signaled arthritis of at least 1 joint, approximately half (n = 43; 38.0%) of this group did not have arthritis (false positive). 2 patients (1.8%) did not signal arthritis, while there was (false negative).
Of the 113 assessments made by parents, 84 (74,3%) signaled arthritis in at least one joint, 42 (37.1%) of these were false-positive. There was one false negative assessment.
When assessment of parent and patient were combined, with the worst (i.e. yellow or red) being the valid one, in 90 (79.6%) out of 113 cases arthritis was signaled. Out of these, 48 (53.3) were false-positive. Combining assessments led to one false negative, which is one less when compared to the patient.
Sensitivity, which measures the proportion of positives which are correctly identified as such (Cat. 4/Cat. 2 + Cat. 4)) is high. Specificity, which measures the proportion of negatives which are correctly identified as such (Cat.1/Cat.1 + Cat.3) is low. Sensitivity and specificity of assessments of child, parent and the combination of groups can be seen in table 1 and figure 2.
Positive predictive value, the percentage of patients or parents that judges that there is arthritis while there is, is low in self and parental assessment.
Negative predictive value is the percentage of the patients that judges there is no arthritis while there is no arthritis and is found to be high in self and parental assessment.

Table2. Assessments by category

		Cat.1 No arthritis=No arthritis	*Cat.2 False negative*	*Cat.3 False positive*	*Cat.4 Arthritis=Arthritis*	*Sens**	*Spec***	*Ppv †*	*Npv††*
Child	Frequency	27	2	43	41	95.4	38.6	49.4	93.1
	Percentage	23.9	1.8	38.1	36.3				
Parent	Frequency	28	1	42	42	97.7	40.0	50.0	96.6
	Percentage	24.8	0.9	37.2	37.2				
Parent+child	Frequency	22	1	48	42	97.7	31.4	46.7	95.7
	Percentage	19.5	0.9	42.5	37.2				

* Sensitivity in %
** Specificity in %
† Positive predive value in %
†† Negative predictive value in %

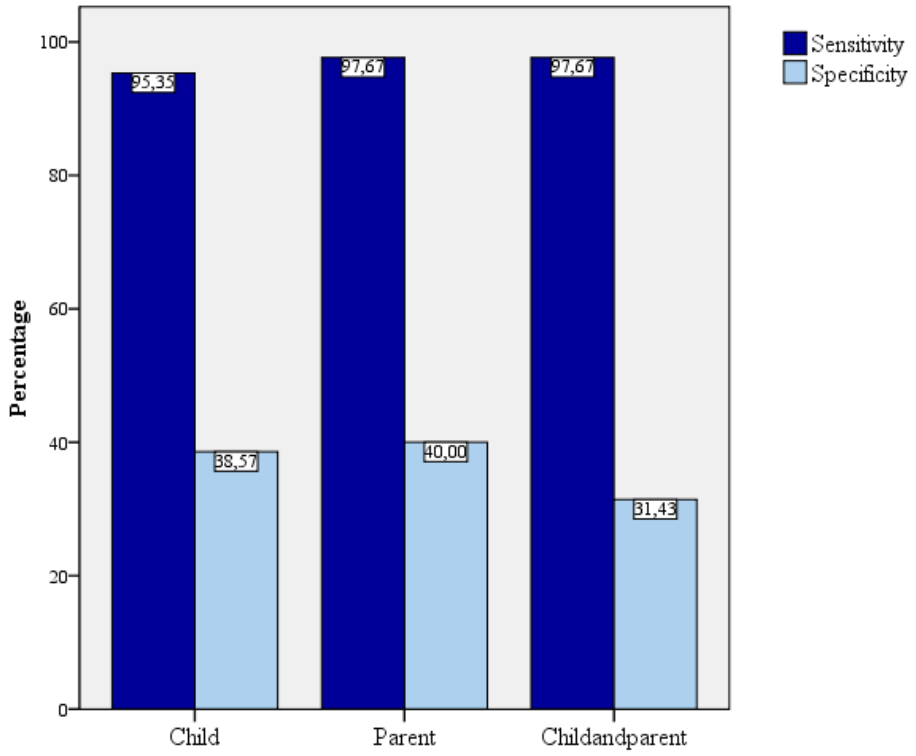


Figure 2. Sensitivity and specificity of child and parental assessment

1) Conclusions

Both in self and parental assessment sensitivity is high and specificity is low. So agreement between physician and patient or parent is low. There are many false positive assessments, thus overrating disease activity. Number of false negative assessments is low, thus patients and parents do not miss arthritis. Combining assessments is of no use, considering it leads to more false-positive assessments, while it only decreases the number of false-negatives with 1(0.9%). This is also reflected in sensitivity and specificity, which are both higher in children and parent. A high negative predictive value and a low positive predictive value leads to the conclusion that a negative assessment (no arthritis) is more reliable than a positive assessment (i.e. arthritis).

Between parent and child a good level of agreement is measured in all 4 categories. In 101 out of 113 cases (89.4%) there is agreement between parent and child. Disagreement is seen in 3 different combinations, summarized in table 3. Two groups contain more than one case. In the first, containing 5 children, child judges correctly that there is no arthritis while the parent judges that there is. In the second group, containing 6 children, the parent judges correctly that there is no arthritis while the child judges that there is (i.e. false positive).
Out of the two false negative assessments made by patients, there was one that was signaled by the parent.

Table 3. Comparing parent assessment versus child assessment based on 4 categories

Child	*Parent* *Cat.1*	*Cat. 2*	*Cat. 3*	*Cat. 4*	**Total**
Cat.1 No arthritis=no arthritis	22▪	0	5*	0	27
Cat. 2 false negative	0	1▪	0	1*	2
Cat. 3 false positive	6*	0	37▪	0	43
Cat. 4 arthritis = arthritis	0	0	0	41▪	41
Total	28	1	42	42	113

▪ agreement
*disagreement

The 41 patients and 42 parents that judged correctly there was arthritis fell in 5 different subcategories, 4a to 4e, depending on localization. This showed that 3 patients (7.3%) and 3 parents (7.1%) marked the wrong joint(s). The largest group (38 of 41; 92.7%) marked the correct joint(s), however 18 of them also marked at least one joint without arthritis, 2 missed at least one joint with arthritis, 12 did both. Only 6 patients and 9 parents localized the correct joint and did not miss or mark any other joints. Again there is little difference between patient and parent.

Table 4. Localization of arthritis by 5 subcategories

	Child		*Parent*		*Parent-child combination*	
	Frequency	*Percentage*	*Frequency*	*Percentage*	*Frequency*	*Percentage*
Subcat 4a: correct location	6	5.3	9	8.0	7	5.9
Subcat 4b wrong location	3	2.7	3	2.7	2	1.7
Subcat 4c correct location+wrong location	18	15.9	17	15.0	23	20.3
Subcat 4d correct location+missed	2	1.8	2	1.8	2	1.7
Subcat 4e correct location+wrong location+missed	12	10.6	11	9.7	8	6.8
Total	41	36.3%	42	37.2%	42	37.2%

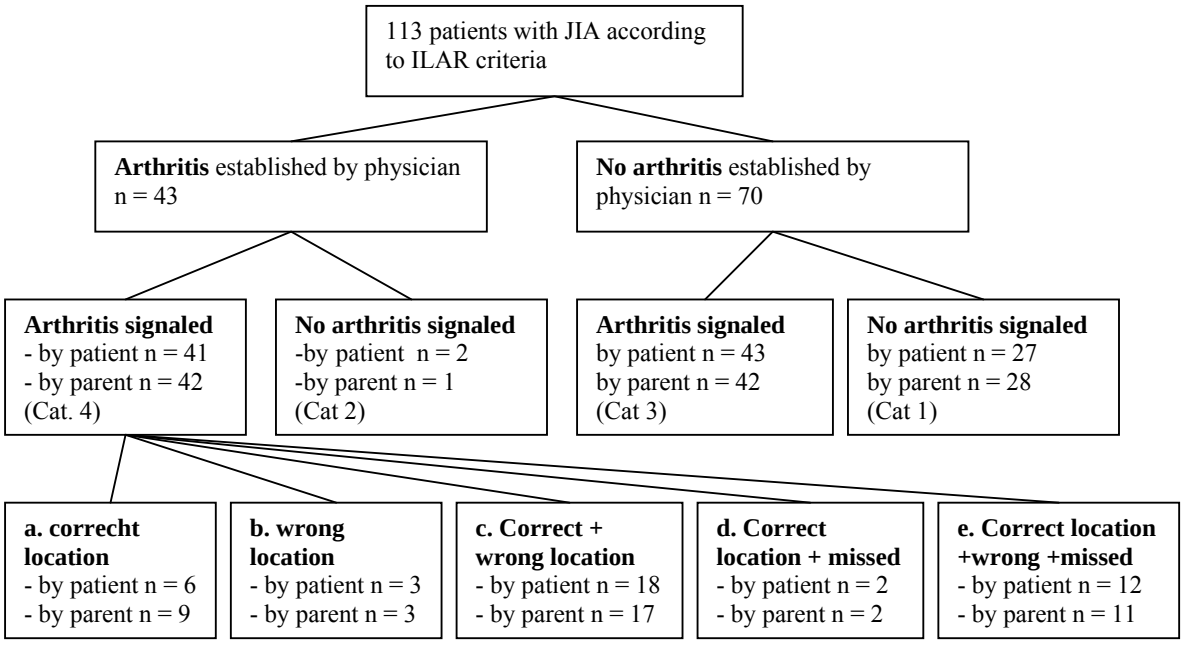


Figure 3. Flow diagram showing the outcome of assessments of 113 patients, made by physicians, parents and patients.

2) Conclusions

Agreement between parent and patient is high with 89.4%. When patients and parents do not agree, one of them unjustly signals arthritis, while the other correctly signals no arthritis. Locating arthritis is good. In the group that signalled arthritis 38 out of 41 marked the correct joint(s). However it also shows that correct assessments are often accompanied by false positive and false negative assessments of other joint(s). Comparing localization of patients and parents, shows that numbers are comparable.
Also considering that combined assessments had no additive value and there was no significant difference in assessments between parent and patient we decided to focus further analysis on the patients' assessments.

3.3 Joint analysis

The assessments of all joints individually and in combinations, forming 4 extremities, were studied next. On this level not the patients but the assessments were labelled with one of the four categories, with the aim to locate which joints or joint groups are most difficult to assess. Outcomes of joint analysis can be seen in table 6.
For clarification this table will be illustrated by the outcome of the left ankle.113 assessments of the left ankle were made. Of these, 78 patients correctly assessed that there was no arthritis (Cat. 1), 3 signalled no arthritis while there was (i.e. false negative; Cat. 2), 20 thought they had arthritis while they did not (i.e. false positive; Cat.3) and 7 patients signalled arthritis in the correctly. Sensitivity varies from 33% to 100%, with a mean of 72.8% and a standard deviation of 18.97. Sensitivity was lowest in the elbow, based on only 3 patients with arthritis. It is highest in the right hip and the left foot (combination), again based on a small number, namely one. Specificity had a less wide range with a minimum of 61.7% a maximum of 96.4%. Mean is 85.4% and the standard deviation is 8.89. Lowest specificity was measured in the left knee. Highest specificity was measured in the left elbow, the joint in which the lowest sensitivity was measured.

Table 6. Self assessment in four categories by joint

	Child					
	Cat 1 No arthritis=no arthritis	*Cat 2 False negative*	*Cat 3 False positive*	*Cat 4 Arthritis=arthritis*	*Sens**	*Spec***
Joint/ joint group						
Ankle L	83	3	20	7	70.0	80.6
Ankle R	78	2	22	11	84.6	78.0
Knee L	58	3	36	16	84.2	61.7
Knee R	57	2	30	24	92.3	65.5
Hip L	95	1	16	1	50.0	85.6
Hip R	96	0	16	1	100	85.7
Cervical spine	97	1	10	5	83.3	90.6
Jaw L	101	1	9	2	66.7	91.8
Jaw R	100	2	8	3	60.0	92.6
Shoulder L	102	1	9	1	50.0	91.9
Shoulder R	102	0	11	0	-	90.2
Elbow L	106	2	4	1	33.3	96.4
Elbow R	100	2	9	2	50.0	91.7
Wrist L + CMC	86	3	19	5	62.5	81.9
Wrist R + CMC	91	1	15	6	85.7	85.8
Small joints hand L	91	3	12	7	70.0	88.3
Small joints hand R	91	1	14	7	87.5	86.7
Small joints foot L	96	0	12	5	100	88.9
Small joints foot R	95	1	13	4	80.0	88.0

* Sensitivity in %** Specificity in %

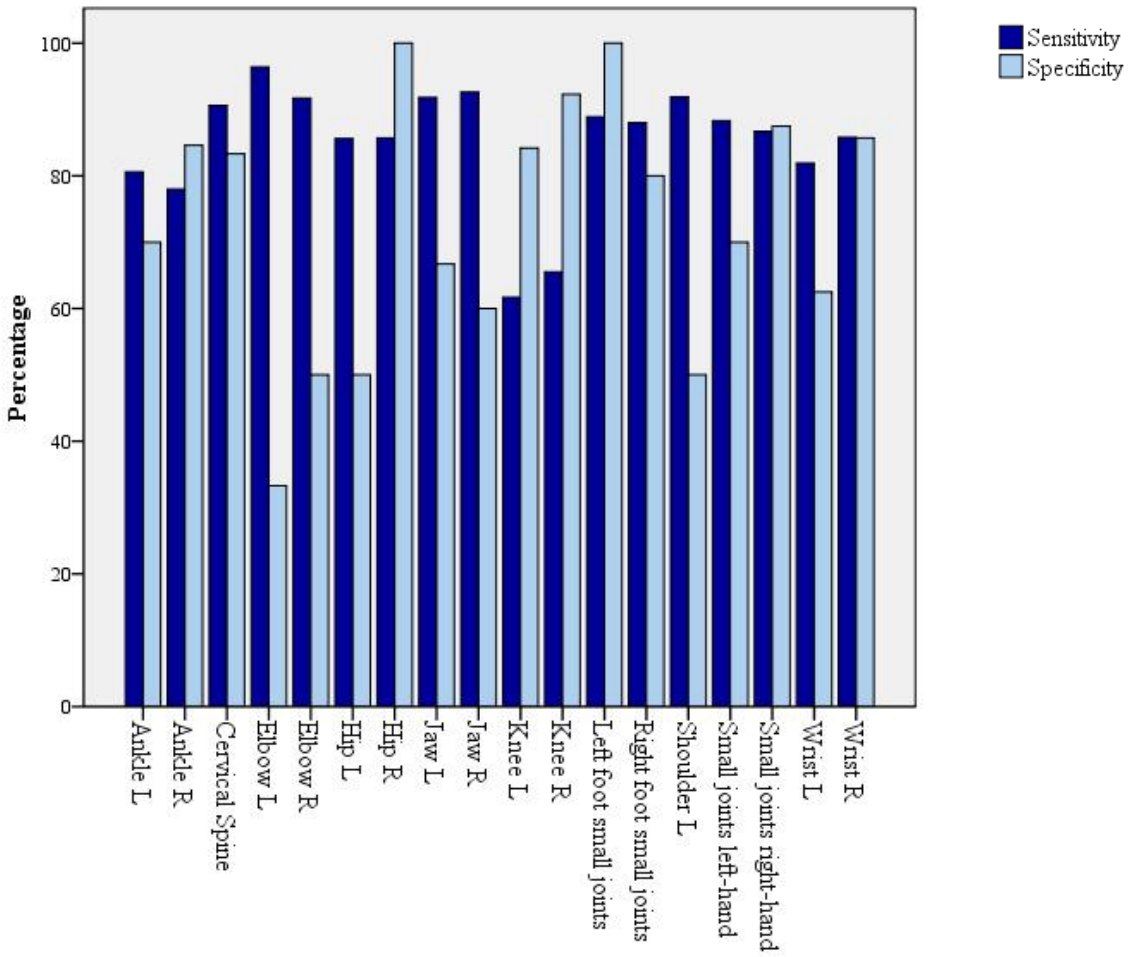


Figure 4. Sensitivity and specificity by joint
*R= right
**L= Left
Because of missing sensitivity, R Shoulder is not displayed.

3) Conclusion

Based on the assessment by joint we can only conclude that based on these numbers nothing can be said of best or worst joint to assess. We could state that arthritis of the L elbow goes unnoticed most, with the detection of only one out of three. However these numbers are very small and thus it is impossible to conclude this, let alone test its significance. However we can also see that joints in which sensitivity and specificity are based on larger numbers (i.e. knees and ankles) a trend towards a high sensitivity and a low specificity is seen, thus again high overrating and low underrating.

Assessment by extremity (seen in table 7) showed that there is a difference in the assessment of right and left. An actual pattern could not be discovered. The right side was usually assessed better (based on 6 values), however there is one exception. The specificity measured in the right arm (74,3%), is lower than the specificity of the left arm (78,3%).

Table 7. Self assessment by extremity

	Child					
	Cat 1	*Cat 2*	*Cat 3*	*Cat 4*		
	No arthritis=no arthritis	*False negative*	*False positive*	*Arthritis=arthritis*	*Sens**	*Spec***
Left arm	76	3	21	13	81.25	78.3
Right arm	75	0	26	12	100	74.3
Left leg	47	3	44	19	86.3	51.6
Right leg	48	1	21	43	97.7	69.6

* Sensitivity in %
** Specificity in %

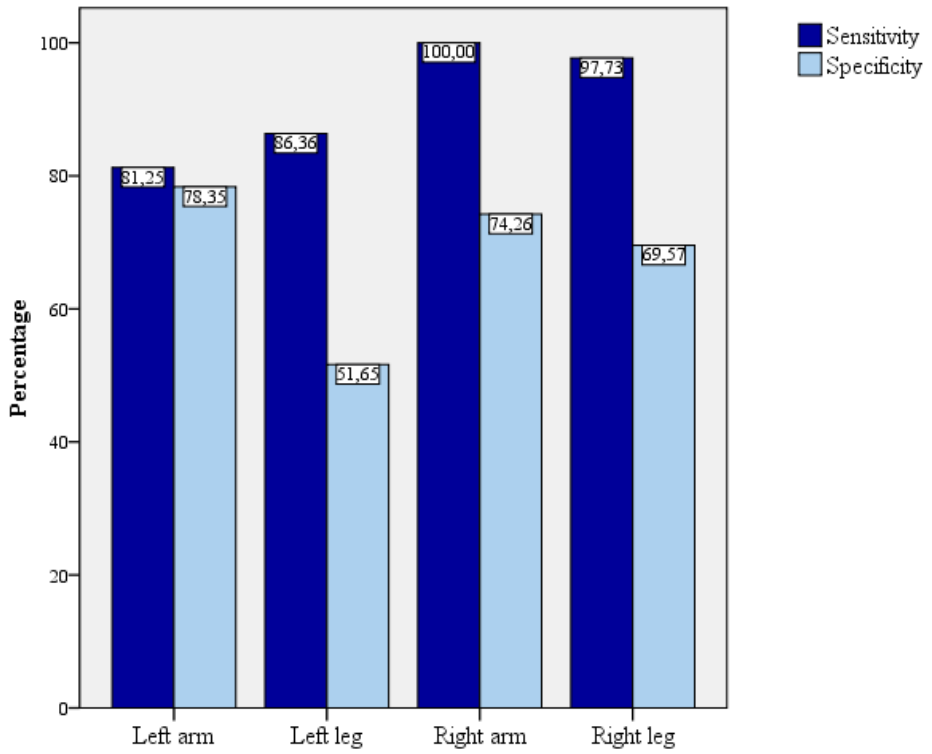


Figure 5. Sensitivity and specificity by extremity

3.4 Predictors

To identify which factors characteristics that could predict discordance between patient and physician we analyzed background variables, such as gender, age, condition, disease duration, VAS pain and CHAQ score. Table 8 summarizes outcomes of descriptive statistics.

Concerning gender, an unequal distribution between males and females and the four categories is seen. In category 1, the true negative assessments, 19.7% of the patients is female, while 32.4% is male. The opposite is seen in Category 3, the false positive assessments, there are more females (43.4%) than males (27%). Of the true positive assessments no relevant difference in gender was seen. The differences found are not significant, Chi- square showed a P- value of 0.291.

Concerning age, the patients that made a true negative assessment are equally distributed between the different age groups. 51.4% of the patients aged between 9 and 12 made a false

positive assessment, while 25.9% and 35.3% of the other age groups, younger than 9 and older than 12 respectively do.
48.1% of the patients younger than 9 and 39.2% of the patients older than 12 are found in the group of true positive assessments, a relevant difference compared to the 22.9% of the patients aged 9-12 it contains. Relevant differences were seen, however no significant difference between the age groups was found.

In relation to condition we can see that the two patients that don't signal arthritis, while it is there, both have oligoarthritis, thus sensitivity is lowest in oligoarthritis. The other patients with an oligo arthritis are equally distributed over the other categories. Of the patients with a polyarthritis 45.5% make false positive assessment, which is high when compared to oligoarthritis (32.6%); ERA (25%) and Systemic (22.2%). There is the exception of undifferentiated arthritis, with 50% of the patients(n=2) that make a false positive assessment, however based on only 2 patients. Differences between the conditions have proved not to be statistically significant, with a P-value of 0.177.

When disease duration is studied, it is seen that the majority, namely 54.5% of patients with short disease duration(less than a year) are found in the group with true positive assessments, compared to 31.9% of the patients with longer disease duration(longer than a year).
42.9 % of the patients with longer disease duration made a false positive assessment, while 18.2 of the patients with shorter disease duration fall in this category. Differences between short and long duration were considered not to be statistically significant, with a P-value of 0.114.

Of the patients with a CHAQ score of zero, 48.9% made a true negative assessment, 31.1% made a false positive assessment and 17.8% made a true positive assessment. 8.9% of the patients with a CHAQ score equal to or lower than one fell in category 1(true negative), the other are distributed between the false positive assessments (44.4%) and the true positive assessments(46.7%). All patients with a CHAQ score higher than one signal arthritis, in 42.1% this is a false positive assessment and in 57.9% this is true positive.
There is a significant difference between categorical CHAQ scores in the assessment of joints (by category). P-value is 0.000.

Concerning the VAS pain scores, we see that both patients with a false negative assessment have a low VAS pain score (<30). The other patients with a VAS pain score lower than 30, are equally distributed over the three other categories, category 1 contains 31.8%, category 3 contains 36.5% and category 4 contains 29.4%. All patients with a VAS score between 30 and 60 signal arthritis, in 54.5% this is incorrect (cat. 3) and in 45.5% this is correct. Of the patients with a VAS score higher than 60, again 100% (n=13) signals arthritis. 38.5% is false positive, while 61.5% is true positive. Differences between the categorical VAS-pain scores are considered statistically significant with a P-value of 0.047.

Table 8. Patient characteristics by category of assessment

Characteristics		*Cat 1 No arthritis = no arthritis* Freq	%	*Cat 2 False negative* Freq	%	*Cat 3 False positive* Freq	%	*Cat 4 Arthritis= Arthritis* Freq	%	*P-value (CHi²)*
Gender	Male	12	32.4	1	2.7	10	27.0	14	37.8	.291
	Female	15	19.7	1	1.3	33	43.4	27	35.5	
Age	<9	6	22.2	1	3.7	7	25.9	13	48.1	.276
	9-12	8	22.9	1	2.9	18	51.4	8	22.9	
	>12	13	23.9	0	-	18	35.3	20	39.2	
Condition	Oligo	13	30.2	2	4.7	14	32.6	14	32.6	.177
	Poly	8	14.5	0	-	25	45.5	22	40.0	
	ERA*	0	-	0	-	1	25.0	3	75.0	
	Systemic	5	55.6	0	-	2	22.2	2	22.2	
	Other**	1	50.0	0	-	1	50	0	-	
Disease duration	=<12 months	6	27.3	0	-	4	18.2	12	54.5	.114
	>=13 months	21	23.1	2	2.2	39	42.9	29	31.9	
CHAQscore†	0	22	48.9	1	2.2	14	31.1	8	17.8	.000
	=<1	4	8.9	0	-	20	44.4	21	46.7	
	>1	0	-	0	-	8	42.1	11	57.9	
VASpain††	<30	27	31.8	2	2.4	31	36.5	25	29.4	.047
	30-60	0	-	0	-	6	54.5	5	45.5	
	>60	0	-	0	-	5	38.5	8	61.5	

* ERA: Enthesitis related arthritis
** Undifferentiated arthritis
† Children's health assessment questionnaire, scores 0-3(categorized in three categories)
†† Visual analogue scale- pain. Measured in mm. scores 0-100(categorized in three categories)
Numbers in bold are found to be remarkable as illustrated in text

Chi-square requires categorical age, disease duration, CHAQ-scores and VAS pain. Without categorization these variables are continuous, as can be seen in table 9. This way an analysis of variances (ANOVA) was obducted. Differences in mean CHAQ scores and VAS-pain scores between the categories are again significant, both with a P-value of 0.000.

Table 9. Analysis of continuing variables by ANOVA

	n=	*Mean*	*Median*	*Minimum*	*Maximum*	*Standard deviation*	P-value (ANOVA)
Age*	113	11.40	12	3	18	3.833	.642
Disease duration**	113	59.08	48	0	192	49.569	.316
CHAQ†	109	0.41	.125	0	2.375	0.516	.000
VAS-pain††	109	20.486	10	0	93	25.68	.000

*Age in years
**Disease duration in months
† Children's health assessment questionnaire scores 0-3
†† Visual analogue scale-pain in mm.

4) Conclusion

Exploration of patient characteristics and the categorical assessment shows many things. A previous find was that, of the patients that don't assess their joints well, most of them overrate. This makes category 3, the false positive assessments, most interesting.

We stated that a difference of 10% between groups and categories was clinically relevant, because of small numbers this could not be tested statistically.

There was no significant difference between males and females, age groups, condition or between shorter and longer disease duration. The difference in gender is clinically relevant, with16.4% more girls that overrate. However difference between boys and girls over the four categories was not statistically significant. Second longer disease duration shows that 18 % more false positive assessments are made, compared to short disease duration. Third, young children seem to be better assessors, with less false positive and more true-positive assessments.

Finally, a significant difference was found between the CHAQ and VAS pain scores, both measured in categories and as continuing variables. All patients with high CHAQ-and VAS pain scores signal arthritis, half of those is false-positive. Thus a high and CHAQ and VAS-score is predictive of signaling arthritis, regardless of being correct or not.

4 Discussion

In this cross-sectional exploratory study we investigated the validity of joint assessment, administered by parent and patient. This was first studied on patient level, then on joint level. We also explored which patient characteristics were of predictive value.
We found, in accord with our hypothesis that the validity of self and parental assessment of joints is low. While 84 patients (74.3%) and also 84 parents signal arthritis, in only half of these cases arthritis is established by the physician. Only two patients (1.8%) and one parent (0.9%) missed arthritis. Thus, sensitivity is high, while specificity is low.
Agreement between parent and patient was found in 101 cases (89.4%), which we considered high. Combining assessments of parent and patient lowered specificity and did not change sensitivity. Patients are able to locate arthritis, but this is often accompanied by overrating of other joints.
Being exploratory, we could not beforehand predict the numbers of discrepancy between patient, parent and physician. But, based on clinical experience we had expected fewer patients to overrate and more patients to underrate. Thus, we underestimated the amount of overrating, possibly because missing arthritis is more impressing to the physician than overrating is. Remarkable is that only two patients contacted the physician because of signalling arthritis. This indicates that a lot of overrating could be explained by insecurity or lack of knowledge about arthritis, unawareness of long term consequences, or fear for (therapeutic) decisions.
In accord with our results, though not actually comparable, are the results of studies of self assessment in rheumatoid arthritis (RA). In these studies the validity of self administered joint assessment is low.(40) A recent study in RA showed that the outcome of the tender joint count was, in accord with our findings, overrated by the patients. However the physician detected more swollen joints. (42)
We expected to see more patients missing arthritis, thus we overestimated underrating. This could be explained as previously said, as a misinterpretation of the physician, but it could also be a sign of denial. Signalling arthritis has diagnostical, therapeutical and sometimes physical consequences. This could also explain that of the patients that signalled arthritis only two contacted the physician.

We had expected a better level of agreement between parent and patient than between physician and parent or patient. In accord with some previous studies in JIA, comparing patients' and parents' judgement of different domains of disease activity, perceived quality of life and functional ability, a good level of agreement was found. However in other studies, comparing judgement of pain and functional ability and rating pain intensity, discordance is found, with parents overrating. (18,29)
When in parents and physicians are compared, concerning functional ability again overrating (i.e. a decreased functional ability) is seen on behalf of the parent.(31) A good level of agreement is found when functional ability is studied.(19)
To explain overrating of both parents and patients in this study leads to speculations, because this has not been studied before in JIA. Moreover, in this study we did not ask patients why they marked joints the way they did. In a previous study parent overrating of functional ability was associated with greater intensity of pain and worse CHAQ score.(19) This is in accord with our findings in which patients with a high CHAQ and VAS pain score more often signal arthritis, false positive (i.e. overrating) or true positive (i.e. correctly signalling arthritis). Indicating that signalling pain is mistaken for signalling arthritis and that evident arthritis is accompanied by more pain than is previously acknowledged.

It is known that underrating has markedly physical consequences.(13) We assume overrating has psychological as well as physical effects. Perceiving sickness while it is not there, could be part of the explanation why patients with JIA have lower levels of physical activity and physical fitness.(34,45)

An important limitation to this study is that we decided to share joints marked with yellow (i.e. doubt) under the term 'arthritis'. To defend this choice, a joint is inflamed, or it is not inflamed. Doubt means that there are clues, which can be swelling, disability, but also pain, that something is wrong with the joint. However a clue could also be that patients or parents have never been convinced that a joint in remission was completely 'free of arthritis'. Important in this matter is that 'doubt' is only stated by physicians in a very few cases, and were al scored in joints difficult to assess, like sacro-iliacal joints, the spine and the jaw joints. The second limitation is in close relation to the first. We included all types of JIA, collected within three months, we did not however relate the joints to disease history. We can only assume that for a patient with oligoarthritis it is easier to assess joints that have never been inflamed than for patients with a polyarthritis. On the other hand we would also assume that patients with a polyarthritis, who usually visit the rheumatologist more often, have increased knowledge or awareness of disease (in)activity. It could also be that patients with polyarthritis have more insecurities and doubts about a joint, because of being more aware of the disease. Furthermore, patients weren't asked why they marked joints yellow or red. Concerning this, we can only speculate. In this study we did not make a separation between swollen and tender joints, which is always done in the studies of self assessment in adults. Clinically we have seen that patients often misjudge pain for arthritis. This could possibly lead to higher overrating. We chose consciously to do this, because the initial intention of this study was to assess whether patients can signal arthritis. Giving further information about arthritis would be an intervention. Though in the instructions (verbal and on paper) we stated this study was about joints with inflammation, and not about pain. Finally, we considered the physician's assessment the gold standard. According to some studies this is not ideal, because of inter-rater variability.(15) However, the four pediatric rheumatologists in this study were all educated in the same institution and worked together for a profound time.

In the analyses of self administered assessment of individual joints and joint groups the same pattern is seen as in patient analysis. Sensitivity is high and specificity is low. Though sensitivity is now based on a small number of cases and shows a wide range, specificity is based on larger numbers, with a less wide variation. There is profound overrating of joints. Which joint is assessed best or worst cannot be stated after this investigation. Combining the joints into extremities did not improve the validity of joint assessment, with again a high sensitivity and a low specificity. However by doing this, we found indications that it is more difficult to assess the left side than the right side.
We had expected that joints that are easy to assess would have a higher specificity and sensitivity. We assumed that these joints would include ankles and knees, considering they are large, hydrops is very much visible and functional impairment causes disability soon. However these joints were overrated most, with 20 false positive assessments of the left ankle, 22 of the right ankle 36 of the left knee and 30 of the right knee.
This could be explained by some the above stated limitations. Including all types of JIA, leads to an increase of particular affected joints, being inflamed or having a history of being inflamed. These particular joints are the knees and ankles.(2) Not relating this to the joint's history (i.e. if the joint had ever been inflamed) and treating doubt as arthritis masks the difference between doubt and arthritis.

We did not expect to find a difference between assessments of the extremities on the right and left side. Our results show, however, not conclusively, that the right side is assessed slightly better. This could be explained by dominance, which we did not record. Another limitation to this study is that the low level of presence of arthritis in most of our patients may have limited the generalizability of this study.

Exploration of patient characteristics and the distribution between the 4 categories shows many things. There was no significant difference between males and females, age groups, condition or between shorter and longer disease duration. However clinically relevant differences were seen. In gender 16.4% girls overrate compared to boys, longer disease duration (>1year) shows that 18 % more false positive assessments are made, young children (<9years) seem to be better assessors, with less false positive and more true-positive assessments.
A significant difference was found between the CHAQ scores, measured as categorical and as a continuing variable. This is corresponding with the VAS pain scores, in which the difference was also significant as categorical and as continuing variable. All patients with high CHAQ and VAS pain scores signal arthritis, half of these is false-positive. Thus a high CHAQ and VAS-pain score is predictive of signaling arthritis, regardless of being correct or not. We expected an equal distribution of boys and girls, of different age groups and of type of condition. In contrast with the outcome we expected patients with disease duration longer than one year to be better assessors. We assume that the finding that patients with a longer disease duration make more false positive assessments is explained by the fact that most patients with a shorter disease duration and that do have active arthritis have not been in remission yet, and there must be greater awareness of disease activity, recently diagnosed.
In accord with our expectations high CHAQ scores (=>1) are found in patients that signal arthritis. Patients with active arthritis have functional impairment. However it seems that patients with a false positive assessment also experience boundaries in activities of daily living. We think that signaling arthritis leads to a high CHAQ-score, which will have its consequences for long term outcome.
In accord with our hypothesis, patients VAS-pain scores above 30 all signal arthritis. Previous studies have shown that patients with JIA have lower pain thresholds (PT), when compared to a reference group. In patients with active disease the PT is even lower. This partly explains the high VAS-scores in the false positive assessments and the even higher VAS-scores in the true positive assessments.(23) It is reported that psychosocial factors are better predictors of pain rating than disease activity is in JIA.(23) These results lead to speculations regarding underrating and under treating pain in children with JIA, which is a current subject in paediatric rheumatology.(23)
There are a few limitations to the exploration of background variables. First, the greater prevalence of females in this cohort might affect the generalizability of these results. However our patients represent a consecutive sampling of our clinic population and are likely representative of the patients seen in most tertiary pediatric rheumatology centers. On this level we studied patients outcome in relation to their assessments, thus no comparison between patient and parent could be made. Furthermore, categorisation of VAS-pain and CHAQ score were based on clinical experience and not on previous findings in literature. Finally number of patients when distributed over 4 categories and at least 2 other variables are too small to statistically test between the categories.

The eventual goal is to make a self or parent administered joint assessment part of the clinical evaluation of JIA. Based on these results we can conclude that the current joint assessment is not valid. A valid joint count would be beneficial for both clinical care and research.

We think education will improve validity. And we assume it will have some other benefits as well. It will increase awareness of disease, which will lead to better self management and eventually an improved perceived quality of life. For parents of patients with JIA it is reported that understanding the medical situation is most useful in improving coping. We think the educational program should focus on patients and parents. Before and after the education, validity of joint assessment should be measured. It would be useful to also measure level of physical activity, perceived quality of life and CHAQ andVAS-pain score before and after education. In this same study parental coping is associated with perceived quality of life of the patient.(9) Educational levels of patients and parents should be recorded. To assess whether there is a true difference between right and left, dominant side should also be recorded. For joint assessment the mannequin should be implemented again. Whether three colours should be used again is a point of discussion. Using green for joints without arthritis seems redundant, but it makes patients assess and think about every joint. 69 joints appeared to be too much in previous studies in RA, a 28 joint count has proven to be just as accurate. The difference made between swollen and tender joints as in adults could be applied as well, another possibility is to ask for clarification when a joint is marked yellow or red. The joint assessment by parent, patient and physician can be expanded by adding ultrasonography. Predictive values should also be studied, in larger amounts of patients. The most effective method of education should be studied. Different age groups and parents should be approached differently.

Conclusively, there is a profound group that does not assess their joints well, most of them overrate, only a few underrate. Locating arthritis is good, but is often accompanied by overrating of other joints. Signalling arthritis did not lead to contacting the physician. This leads to the conclusion that while patients can not only not assess joints, we assume that they are unaware of symptoms of arthritis and its long-term consequences, how and when it should be treated and when to contact the physician.
Because there is no cure for JIA, and because no treatments have been demonstrated to result in prolonged periods of disease remission, achieving self management is essential, with a central role for accurate self assessment. This might eventually improve effects, consequences and the experience of JIA concerning physical, educational, functional and psychosocial burdens and the potential for ongoing joint damage. First and most important is that based on these results, education is very much needed.

Afterword and acknowledgements

If it was a lucky coincidence that this turned out to be the project for my master thesis I don't know. However it was a coincidence and I do feel lucky. For a person that has never been that much of a researcher, this project seemed ideal. I wanted to do something that was closely related to clinical practice, that would inspire me to learn something about statistics, starting with a blank canvas was essential and it had to have something to do with pediatrics. This study was basic in its formula, I had a very good reference about the rheumatologist I was about to work with and, besides the idea and a hypothesis, there was just nothing there yet.

Piles of paper, empty felt-tips, the owner of a huge database, a growing complexity in analyses and the realization of the importance of what we were doing could, retrospectively have been the moment for a breakdown. It would have been, if it weren't for all the help we got. It has never been just me doing this research. Support on different levels, but each of them indispensable. The best advice, inspiration, enthusiasm and lots of chocolate and coffee led to a product of which I am proud to present as my master thesis. But that's not most important, I sincerely hope that it will eventually lead to an educational program that will enhance the quality of life of a group of patients that really can use the improvement. Of course I'm a little biased, but as far as I have seen, the room for improvement is there.

Realizing that this is only a master thesis (that has not even been approved of yet) and no receipt of an Oscar or the afterword of a dissertation it would be a little presumptuous to name names.

Also realizing that not everyone knows their importance and that I really do need to say thanks. Wineke, for being the most important person in the realization of this project. Thank you for taking the risk of working with me. For your knowledge, your enthusiasm and support. For being hardworking, honest and everything else Jelmer promised. Eric for teaching me all I know about statistics, the almost therapeutic sessions, your huge share in the analyses and your trust in me. Jelte for providing both wisdom and coffee. Jelmer, for contacting Wineke. Brigitte, Ina, Otto, Klaasje, Ennelien, Elsbeth, Hannie, Liesbeth and Peter for help, support, tolerating me and the delays I caused, for sharing desks, stories, music, cookies, cucumbers, chocolate and coffee. Binnur, Alberta, Annemarie, dr. Wulffraat, dr. Prakken and dr. Vastert for all the help and making me feel at home in Utrecht. Dr. Leijsma for the inclusion of adolescents. Last but not least, thanks to all 113 patients and parents, without your help and enthusiasm, this would never have been possible.

References

(1) Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J.Rheumatol. 2004 Feb;31(2):390-392.

(2) Woo P, Laxer R.M., Sherry ED. Pediatric Rheumatology in clinical practice. 1st ed. London: Springer; 2007.

(3) Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007 Mar 3;369(9563):767-778.

(4) Martini A, Lovell DJ. Juvenile idiopathic arthritis: state of the art and future perspectives. Ann.Rheum.Dis. 2010 Jul;69(7):1260-1263.

(5) Wallace CA, Huang B, Bandeira M, Ravelli A, Giannini EH. Patterns of clinical remission in select categories of juvenile idiopathic arthritis. Arthritis Rheum. 2005 Nov;52(11):3554-3562.

(6) Magnani A, Pistorio A, Magni-Manzoni S, Falcone A, Lombardini G, Bandeira M, et al. Achievement of a state of inactive disease at least once in the first 5 years predicts better outcome of patients with polyarticular juvenile idiopathic arthritis. J.Rheumatol. 2009 Mar;36(3):628-634.

(7) van der Kooij SM, de Vries-Bouwstra JK, Goekoop-Ruiterman YP, Ewals JA, Han KH, Hazes JM, et al. Patient-reported outcomes in a randomized trial comparing four different treatment strategies in recent-onset rheumatoid arthritis. Arthritis Rheum. 2009 Jan 15;61(1):4-12.

(8) Cavallo S, Feldman DE, Swaine B, Meshefedjian G. Is parental coping associated with the level of function in children with physical disabilities? Child Care Health Dev. 2009 Jan;35(1):33-40.

(9) Cavallo S, Feldman DE, Swaine B, Meshefedjian G, Malleson PN, Duffy CM. Is parental coping associated with quality of life in juvenile idiopathic arthritis? Pediatr.Rheumatol.Online J. 2009 Mar 11;7:7.

(10) Ravelli A, Martini A. Early predictors of outcome in juvenile idiopathic arthritis. Clin.Exp.Rheumatol. 2003 Sep-Oct;21(5 Suppl 31):S89-93.

(11) Ravelli A. Toward an understanding of the long-term outcome of juvenile idiopathic arthritis. Clin.Exp.Rheumatol. 2004 May-Jun;22(3):271-275.

(12) Minden K. Adult outcomes of patients with juvenile idiopathic arthritis. Horm.Res. 2009 Nov;72 Suppl 1:20-25.

(13) Minden K, Niewerth M, Listing J, Biedermann T, Bollow M, Schontube M, et al. Long-term outcome in patients with juvenile idiopathic arthritis. Arthritis Rheum. 2002 Sep;46(9):2392-2401.

(14) Pincus T, Sokka T. Quantitative measures for assessing rheumatoid arthritis in clinical trials and clinical care. Best Pract.Res.Clin.Rheumatol. 2003 Oct;17(5):753-781.

(15) Scott DL, Choy EH, Greeves A, Isenberg D, Kassinor D, Rankin E, et al. Standardising joint assessment in rheumatoid arthritis. Clin.Rheumatol. 1996 Nov;15(6):579-582.

(16) April KT, Feldman DE, Platt RW, Duffy CM. Comparison between Children with Juvenile Idiopathic Arthritis (JIA) and their parents concerning perceived Quality of Life. Qual.Life Res. 2006 May;15(4):655-661.

(17) Consolaro A, Vitale R, Pistorio A, Lattanzi B, Ruperto N, Malattia C, et al. Physicians' and parents' ratings of inactive disease are frequently discordant in juvenile idiopathic arthritis. J.Rheumatol. 2007 Aug;34(8):1773-1776.

(18) Palermo TM, Zebracki K, Cox S, Newman AJ, Singer NG. Juvenile idiopathic arthritis: parent-child discrepancy on reports of pain and disability. J.Rheumatol. 2004 Sep;31(9):1840-1846.

(19) Palmisani E, Solari N, Pistorio A, Ruperto N, Malattia C, Viola S, et al. Agreement between physicians and parents in rating functional ability of children with juvenile idiopathic arthritis. Pediatr.Rheumatol.Online J. 2007 Dec 11;5:23.

(20) Sawyer MG, Carbone JA, Whitham JN, Roberton DM, Taplin JE, Varni JW, et al. The relationship between health-related quality of life, pain, and coping strategies in juvenile arthritis--a one year prospective study. Qual.Life Res. 2005 Aug;14(6):1585-1598.

(21) Ringold S, Seidel KD, Koepsell TD, Wallace CA. Inactive disease in polyarticular juvenile idiopathic arthritis: current patterns and associations. Rheumatology (Oxford) 2009 Aug;48(8):972-977.

(22) Ringold S, Wallace CA, Rivara FP. Health-related quality of life, physical function, fatigue, and disease activity in children with established polyarticular juvenile idiopathic arthritis. J.Rheumatol. 2009 Jun;36(6):1330-1336.

(23) Thastum M, Zachariae R, Herlin T. Pain experience and pain coping strategies in children with juvenile idiopathic arthritis. J.Rheumatol. 2001 May;28(5):1091-1098.

(24) Stanford EA, Chambers CT, Craig KD. The role of developmental factors in predicting young children's use of a self-report scale for pain. Pain 2006 Jan;120(1-2):16-23.

(25) Stinson JN, Kavanagh T, Yamada J, Gill N, Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. Pain 2006 Nov;125(1-2):143-157.

(26) Kuis W, Heijnen CJ, Sinnema G, Kavelaars A, van der Net J, Helders PJ. Pain in childhood rheumatic arthritis. Baillieres Clin.Rheumatol. 1998 May;12(2):229-244.

(27) Stanford EA, Chambers CT, Craig KD. The role of developmental factors in predicting young children's use of a self-report scale for pain. Pain 2006 Jan;120(1-2):16-23.

(28) Cohen LL, Lemanek K, Blount RL, Dahlquist LM, Lim CS, Palermo TM, et al. Evidence-based assessment of pediatric pain. J.Pediatr.Psychol. 2008 Oct;33(9):939-55; discussion 956-7.

(29) Garcia-Munitis P, Bandeira M, Pistorio A, Magni-Manzoni S, Ruperto N, Schivo A, et al. Level of agreement between children, parents, and physicians in rating pain intensity in juvenile idiopathic arthritis. Arthritis Rheum. 2006 Apr 15;55(2):177-183.

(30) Hogeweg JA, Kuis W, Huygen AC, de Jong-de vos van Steenwijk,C., Bernards AT, Oostendorp RA, et al. The pain threshold in juvenile chronic arthritis. Br.J.Rheumatol. 1995 Jan;34(1):61-67.

(31) Ravelli A, Viola S, Migliavacca D, Pistorio A, Ruperto N, Martini A. Discordance between proxy-reported and observed assessment of functional ability of children with juvenile idiopathic arthritis. Rheumatology (Oxford) 2001 Aug;40(8):914-919.

(32) Wulffraat N, van der Net JJ, Ruperto N, Kamphuis S, Prakken BJ, Ten Cate R, et al. The Dutch version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ). Clin.Exp.Rheumatol. 2001 Jul-Aug;19(4 Suppl 23):S111-5.

(33) Long AR, Rouster-Stevens KA. The role of exercise therapy in the management of juvenile idiopathic arthritis. Curr.Opin.Rheumatol. 2010 Mar;22(2):213-217.

(34) Lelieveld OT, Armbrust W, van Leeuwen MA, Duppen N, Geertzen JH, Sauer PJ, et al. Physical activity in adolescents with juvenile idiopathic arthritis. Arthritis Rheum. 2008 Oct 15;59(10):1379-1384.

(35) Singh-Grewal D, Schneiderman-Walker J, Wright V, Bar-Or O, Beyene J, Selvadurai H, et al. The effects of vigorous exercise training on physical function in children with arthritis: a randomized, controlled, single-blinded trial. Arthritis Rheum. 2007 Oct 15;57(7):1202-1210.

(36) Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. Psychol.Bull. 1987 Mar;101(2):213-232.

(37) van der Heide A, Jacobs JW, Dinant HJ, Bijlsma JW. The impact of endpoint measures in rheumatoid arthritis clinical trials. Semin.Arthritis Rheum. 1992 Apr;21(5):287-294.

(38) Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. Ann.Intern.Med. 1994 Jan 1;120(1):26-34.

(39) Prevoo ML, Kuper IH, van't Hof MA, van Leeuwen MA, van de Putte LB, van Riel PL. Validity and reproducibility of self-administered joint counts. A prospective longitudinal followup study in patients with rheumatoid arthritis. J.Rheumatol. 1996 May;23(5):841-845.

(40) Barton JL, Criswell LA, Kaiser R, Chen YH, Schillinger D. Systematic review and metaanalysis of patient self-report versus trained assessor joint counts in rheumatoid arthritis. J.Rheumatol. 2009 Dec;36(12):2635-2641.

(41) Spoorenberg A, van der Heijde D, Dougados M, de Vlam K, Mielants H, van de Tempel H, et al. Reliability of self assessed joint counts in ankylosing spondylitis. Ann.Rheum.Dis. 2002 Sep;61(9):799-803.

(42) Cheung PP, Ruyssen-Witrand A, Gossec L, Paternotte S, Le Bourlout C, Mazieres M, et al. Reliability of patient self-evaluation of swollen and tender joints in rheumatoid arthritis: A comparison study with ultrasonography, physician, and nurse assessments. Arthritis Care.Res.(Hoboken) 2010 Aug;62(8):1112-1119.

(43) Osborne RH, Wilson T, Lorig KR, McColl GJ. Does self-management lead to sustainable health benefits in people with arthritis? A 2-year transition study of 452 Australians. J.Rheumatol. 2007 May;34(5):1112-1117.

(44) Lorig K, Gonzalez VM, Ritter P. Community-based Spanish language arthritis education program: a randomized trial. Med.Care 1999 Sep;37(9):957-963.

(45) Takken T, Van Brussel M, Engelbert RH, Van Der Net J, Kuis W, Helders PJ. Exercise therapy in juvenile idiopathic arthritis: a Cochrane Review. Eur.J.Phys.Rehabil.Med. 2008 Sep;44(3):287-297.